УДК 519.7

Олександр Галкін
Київський національний університет імені Тараса Шевченка, факультет кібернетики, Україна

КОЕФІЦІЄНТИ МАСШТАБУВАННЯ ТА ГРЕБЕНЕВА РЕГРЕСІЯ

OleksandrGalkin

SCALING FACTORS AND RIDGE REGRESSION

Автоматичне знаходження коефіцієнтів масштабування відіграє важливу роль в інтелектуальному аналізі даних, оскільки вхідні компоненти можуть бути як певними мірами різного характеру, так і виконувати відповідне зважування вхідних характеристик. Ґрунтуючись на реалізації автоматичного вибору коефіцієнтів масштабування для опорно-векторних машин в контексті класифікації, ми розглядаємо подібну методологію, але в контексті регресії. Зокрема, ми побачимо, як за допомогою градієнтного спуску знаходити коефіцієнти масштабування, які мінімізують помилку пропуску алгоритму гребеневої регресії.

Гребенева регресія. Розглянемо множину функцій

$$f(x,\alpha)=\sum_{i=1}^{p}\alpha_i\varphi_i(x),$$

де φ_i є базисними функціями. У методах ядра, зазвичай $p=n$ та $\varphi_i=K(x_i,\cdot)$.

Алгоритм гребеневої регресії полягає в зведені до мінімуму наступного функціоналу:

$$R_{emp}(\alpha)=\frac{1}{n}\sum_{i=1}^{n}(y_i-f(x_i,\alpha))^2+\gamma\|\alpha\|^2,$$

де γ є фіксованою позитивної постійною величиною, що називається *параметром регуляризації*. Мінімум задається вектором коефіцієнтів

$$\alpha^0=(K^TK+\gamma I)^{-1}K^TY,$$

де

$$Y=(y_1,...,y_n)^T,$$

а K є матрицею з елементами

$$K_{ij}=\varphi_j(x_i), i=1,...,n, j=1,...,p.$$

Помилка пропуску. Корисною властивістю алгоритму гребеневої регресії є те, що його помилка пропуску має замкнуту форму. Дійсно, позначаючи

$$A_\gamma=K^TK+\gamma I,$$

помилка, зумовлена процедурою пропуску, є:

$$T_{np}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i-k_i^TA_\gamma^{-1}K^TY}{1-k_i^TA_\gamma^{-1}k_i}\right)^2,$$

де

$$k_i=(\varphi_1(x_i),...,\varphi_p(x_i))^T.$$

Чисельник визначає відхилення значень оцінок гребеневої регресії в навчальних точках від істинних значень, тобто навчальної помилки. Знаменник визначає корекцію до цих оцінок для процедури пропуску.

Оптимізація коефіцієнтів масштабування. Тепер припустимо, що $\varphi_i=K(x_i,\cdot)$, де K є ядром радіальної базисної функції із діагональною коваріаційною матрицею, що задається рівнянням

$$K(x,y) = \exp\left(-\sum_i \frac{(x_i - y_i)^2}{2\sigma_i^2} \right).$$

Таким чином, гіперпараметрами алгоритмів гребеневої регресії є $\gamma, \sigma_1, ..., \sigma_d$, які можуть бути оптимізовані з використанням градієнтному спуску на (2). Дійсно, аналітично можуть бути проведені обчислення

$$\frac{\partial T_{np}}{\partial \sigma_p} \text{ та } \frac{\partial T_{np}}{\partial \gamma}.$$

Для того, щоб звести до мінімуму помилку пропуску T_{np}, ми використали оптимізаційний інструментарій Matlab, а для того, щоб уникнути додавання обмежень $\sigma_p \geq 0$ та $\gamma \geq 0$, була проведена оптимізація на $\log \sigma_p$ та $\log \gamma$.

Підсумовуючи викладений матеріал, зазначимо, що ґрунтуючись на реалізації автоматичного вибору коефіцієнтів масштабування для опорно-векторних машин в контексті класифікації, була розглянута схожа методологія, однак в контексті регресії. Було встановлено, як за допомогою градієнтного спуску знаходити коефіцієнти масштабування, які мінімізують помилку пропуску алгоритму гребеневої регресії.

Література

1. Goutte C. and Larsen J., “Adaptive metric kernel regression”, In Neural Networks for Signal Processing VIII, Piscataway, New Jersey, 1998.